One Objective to Align Them All: Variational Free Energy and KL Divergence from Brains to Logostics

Douglas C. Youvan

doug@youvan.com

www.youvan.ai

November 6, 2025

Why do variational free energy (VFE) and the Kullback–Leibler (KL) divergence keep showing up together—across neuroscience, robotics, machine learning, and control—and why are they central to our Logostic q(t) → τ(t) program? The short answer is that VFE/KL provide a single, principled scalar that is simultaneously a bound on surprise (ELBO), a code-length surrogate (MDL), a Lyapunov-like contraction index, and a geometry (via Fisher information) for taking invariant steps toward alignment. In Friston's active inference, they knit perception, learning, and action into one update rule; in VAEs and modern variational inference, they train latent-variable models; in control-as-inference and KL-control, they reshape policies by penalizing divergence from passive dynamics; in robotics, they become a generalized gradient descent on world-model error. This paper surveys those threads and then refactors them into a coherent Logostics view: use VFE/KL as the global contraction functional that drives q toward τ while letting "membranes" (Markov partitions) localize alignment pressure. We compare instantaneous vs expected free energy, accuracy–complexity decompositions, and natural-gradient flows; we show how to instrument per-membrane KLs, escape-time maps, and contraction indices; and we sketch when to switch to Wasserstein or α-divergences. The aim is practical: a portable, domain-agnostic recipe for building iterative generators that contract q(t) toward τ(t) while deliberately surfing boundaries where insight emerges.

Keywords: variational free energy, KL divergence, ELBO, active inference, control as inference, KL control, variational inference, VAEs, information geometry, Fisher metric, Markov membrane, Logostics, q(t) → τ(t), Lyapunov function, MDL.

A collaboration with GPT-5. 38 pages. CC4.0.

Outline

1. Motivation and Claim
   1.1 Why VFE/KL recur across fields
   1.2 The Logostics thesis: one scalar for distance, code, and control
2. Primer: VFE, KL, and Information Geometry
   2.1 From ELBO to free energy: $F = KL(q \,||\, p_\tau(\cdot|x)) - \log p_\tau(x)$
   2.2 Accuracy–complexity split and bits-back view
   2.3 Fisher metric and natural gradients as the local geometry of KL
3. Where They Co-Emerge (Survey)
   3.1 Neuroscience & active inference (perception, learning, action)
   3.2 Machine learning: VI, VAEs, amortization, reparameterization
   3.3 Control as inference & KL-control; linearly-solvable MDPs
   3.4 Robotics: generalized gradient descent on VFE
   3.5 Computational psychiatry & theoretical biology: organisms minimizing VFE
4. From Survey to Synthesis: Logostics $q(t) \rightarrow \tau(t)$
   4.1 Treat $F[q \,||\, \tau]$ as global contraction functional
   4.2 Markov membranes: localizing alignment via conditional independencies
   4.3 Instantaneous vs expected free energy for policy selection
   4.4 Natural-gradient flows and invariance under reparameterization
5. Design Pattern: Building $q \rightarrow \tau$ Systems
   5.1 Two-timescale routine: inner KL contraction, outer τ retuning
   5.2 Diagnostics: per-membrane KLs, contraction indices, escape-time maps
   5.3 Boundary surfing: controlled excursions at alignment phase flips
6. When KL Is Not Enough
   6.1 α-divergences to trade off mode-seeking vs mode-covering
   6.2 Wasserstein for support mismatch and dynamics in data space
   6.3 Hybrid objectives and schedule switching across training phases
7. Case Studies (Sketches)
   7.1 Active inference agent with expected-F policies
   7.2 VAE with membrane-wise KL budgetting
   7.3 KL-control for safe exploration under model uncertainty

8. Theoretical Notes
   8.1 Conditions for F as a Lyapunov-like function with moving τ
   8.2 Relation to MDL, PAC-Bayes, and stochastic thermodynamics
   8.3 Links to score matching and diffusion objectives

9. Open Problems and Next Steps
   9.1 Membrane discovery and adaptive partitions
   9.2 Multi-agent Logostics and shared τ fields
   9.3 Toward principled boundary metrics for “insight regions”

10. Conclusion
   10.1 Why VFE/KL are the workhorse of alignment
   10.2 A portable recipe for contracting q(t) toward τ(t) across domains

References

1. Motivation and Claim

1.1 Why VFE/KL recur across fields

VFE (variational free energy) and KL divergence keep reappearing because they jointly solve three perennial problems with one instrument:

1. Inference under limited models.
   When the world model p_τ is misspecified or intractable, the ELBO turns negative VFE into a computable surrogate of surprise:
   $-F = \text{ELBO} = E_q[\log p_\tau(x,z)] - E_q[\log q(z|x)]$.
   Maximizing ELBO (minimizing F) yields the best-available posterior q in the model's family and does so with stochastic gradients that scale.

2. Communication and compression.
   The VFE split is an MDL ledger:
   F = complexity - accuracy
   $= KL(q(z|x) \,||\, p_\tau(z)) - E_q[\log p_\tau(x|z)]$.
   Lower F implies shorter expected codelengths. This binds *learning* (fit the world) to *coding* (compress the world) without extra machinery.

3. Dynamics and control.
   The local quadratic form of KL endows parameter space with the Fisher information metric, turning gradient steps on F into *natural* steps that are invariant to reparameterization. In control-as-inference and active inference, this same scalar (instantaneous or expected) schedules actions that reduce future surprise while honoring prior preferences.

Because a single F:

- upper-bounds surprise (statistics),
- upper-bounds description length (information theory),
- and contracts trajectories (dynamical systems),

it becomes the "lowest common denominator" objective in neuroscience (active inference), machine learning (VI/VAEs), robotics (state-estimation-as-inference), and decision-making (KL control). The asymmetry of $KL(q \,||\, p_\tau)$ is a feature: it punishes putting mass where τ says there is none, a crucial safeguard when aligning a learner to a moving truth-manifold.

Operational payoff: VFE/KL give stable scalars, decomposable across conditional-independence boundaries (our "membranes"), with gradients that are tractable, amortizable, and composable—ideal for multi-scale systems that must learn while acting.

1.2 The Logostics thesis: one scalar for distance, code, and control

Thesis. In Logostics, let τ be a generative law over latent z and data x, and $q \equiv q_\theta(z|x)$ be the agent state. Then a single scalar,
$F[q \,||\, \tau] = KL(q(z|x) \,||\, p_\tau(z|x)) - \log p_\tau(x)$,
is sufficient to play three roles simultaneously:

- Distance (alignment metric).
  For a fixed τ, minimizing F is equivalent to minimizing $KL(q \,||\, p_\tau(\cdot|x))$. Under mild regularity and slowly varying $\tau(t)$, F serves as a Lyapunov-like functional: if $\dot{q} \propto -\nabla_q F$, then $F[q_t \,||\, \tau_t]$ decreases on average, contracting $q(t)$ toward $\tau(t)$.

- Code (Occam accounting).
  The complexity–accuracy split of F enforces parsimony: if $KL(q \,||\, p_\tau(z))$ grows, alignment stalls even when likelihood improves. This lets us *budget* alignment pressure across "membranes" (subsystems, sensors, modules) and detect whether misalignment is due to overfitting internals or underfitting externals.

- Control (policy selection).
  The expected free energy, $G = E_{future}[F]$, decomposes into epistemic and pragmatic terms. Minimizing G selects actions that both reduce uncertainty and realize preferred outcomes encoded by τ. Thus $q(t) \rightarrow \tau(t)$ is not only a *learning* dynamic but a *steering* principle.

Design pattern (two timescales).

- Inner loop (inference): update q to reduce instantaneous F holding τ fixed (fast contraction).

- Outer loop (model/goal retuning): update τ (parameters, structure, or priors over outcomes) using held-out evidence and boundary diagnostics (slow reshaping of the attractor).

Boundary calculus (membranes).
Factorizations implied by Markov membranes make F additive: $F = \Sigma_k F_k$. We can therefore localize gradients, assign KL budgets per boundary, and *surf* controlled phase flips (where contraction signs change) to harvest insight without losing global stability.

Testable predictions.

1. Systems that monitor per-membrane KLs and constrain total F will exhibit faster, stabler convergence than ones optimizing raw loss.

2. Natural-gradient descent on F will reduce sensitivity to parameterization and improve alignment under distribution shift.
3. Scheduling exploration by minimizing expected F (not ad hoc bonuses) will yield better sample efficiency in nonstationary tasks.

Claim. Treating VFE/KL as the *one scalar*—distance-to-τ, code-length ledger, and control objective—turns alignment from a slogan into an implementable routine: contract q(t) toward τ(t), keep the code honest, and act to make future alignment easier.

2. Primer: VFE, KL, and Information Geometry

2.1 From ELBO to free energy: F = KL(q || p_τ(·|x)) − log p_τ(x)

Setup. Let τ denote a generative model p_τ(x,z) = p_τ(x|z) p_τ(z). Let q(z|x) be an approximate posterior (the agent's state). Then

ELBO(q; τ) = E_q[log p_τ(x,z)] − E_q[log q(z|x)].

Define the (variational) free energy as

F(q || τ) = − ELBO(q; τ).

Two key identities:

1. Bound on surprise (Jensen step).
   log p_τ(x) = ELBO(q; τ) + KL(q(z|x) || p_τ(z|x))
   ⇒ −log p_τ(x) = F(q || τ) − KL(q || p_τ(·|x))
   ⇒ F(q || τ) = KL(q || p_τ(·|x)) − log p_τ(x).

Thus for fixed τ and data x, minimizing F over q is equivalent to minimizing KL(q || p_τ(z|x)); the gap to negative log evidence is exactly that KL.

2. Variational inference as projection.
   argmin_q F(q || τ) = argmin_q KL(q || p_τ(z|x))
   ⇒ q*(z|x) = p_τ(z|x) if q's family is rich enough.
   With restricted q, F chooses the "closest" posterior in the KL sense.

Interpretations:

- Statistics: −log p_τ(x) is model surprise; F upper-bounds it by a tractable objective.
- Coding: −ELBO is an expected codelength (Sec. 2.2).
- Dynamics: F is a scalar "distance-to-τ" we can descend (Sec. 2.3).

2.2 Accuracy–complexity split and bits-back view

Start from the joint form of ELBO:

$ELBO = E_q[\log p_\tau(x|z)] + E_q[\log p_\tau(z)] - E_q[\log q(z|x)]$.

Rearrange:

$F(q \,||\, \tau) = -ELBO$
$= [\, - E_q \log p_\tau(x|z) \,] + KL(q(z|x) \,||\, p_\tau(z))$
= accuracy term + complexity term.

- Accuracy = $- E_q \log p_\tau(x|z)$: expected negative log-likelihood under q. Lower is better fit to data.
- Complexity = $KL(q(z|x) \,||\, p_\tau(z))$: how far the posterior is from the prior. Lower is simpler, more “Occam”.

Budgeting intuition. If accuracy improves while complexity explodes, F may stall; the split tells you *why* alignment fails (bad likelihood vs overfitting latent beliefs).

Bits-back coding. Map each term to codelengths (in nats or bits):

- An optimal code for z from the prior would cost $-\log p_\tau(z)$.
- If we instead sample $z \sim q(z|x)$, we pay $-\log q(z|x)$ but can “get back” $-\log p_\tau(z)$ bits by reusing the prior code.
- The net overhead is $KL(q \,||\, p_\tau(z)) = E_q[\log q - \log p_\tau(z)]$.

Total expected message length for x under q is

$L(q; \tau) = E_q[\, -\log p_\tau(x|z) \,] + KL(q \,||\, p_\tau(z)) = F(q \,||\, \tau)$.

Therefore: minimizing F is the same as minimizing expected codelength (MDL). This is why VFE/KL unify *learning* and *compression*, and why monitoring the split is a practical diagnostic: you can watch whether you are paying for fit (accuracy) or for belief complexity (complexity).

Membrane additivity. If the model factorizes across a Markov membrane into conditionally independent parts, both accuracy and complexity terms add over factors. That makes F decomposable: perfect for local budgets and per-boundary audits.

2.3 Fisher metric and natural gradients as the local geometry of KL

KL induces a Riemannian geometry on statistical manifolds. Consider a parametric family $p_\theta(y)$. For a small step $d\theta$:

$KL(p_\theta || p_{\theta+d\theta}) = (1/2)\, d\theta^T I(\theta)\, d\theta + o(\|d\theta\|^2)$,

where $I(\theta)$ is the Fisher information matrix,

$I(\theta) = E_{y\sim p_\theta} [(\nabla_\theta \log p_\theta(y)) (\nabla_\theta \log p_\theta(y))^T]$.

Consequences:

1. Local quadratic form. Close to θ, KL behaves like a squared Mahalanobis distance with metric $I(\theta)$. This is the *information geometry* that makes learning steps well-scaled.
2. Natural gradient. Given an objective $L(\theta)$, the steepest descent direction under the KL-induced geometry solves

minimize $d\theta^T \nabla_\theta L(\theta)$ subject to $d\theta^T I(\theta)\, d\theta = \varepsilon$,

yielding the *natural gradient* update

$\theta_{t+1} = \theta_t - \eta\, I(\theta_t)^{-1} \nabla_\theta L(\theta_t)$.

Properties:

- Reparameterization invariance. Unlike vanilla gradients, natural gradients do not depend on how you numerically encode θ; they follow the intrinsic geometry of the model manifold.
- Better conditioning. I^{-1} rescales directions by sensitivity, often accelerating convergence and stabilizing training when likelihood curvature is skewed.

3. Applying to F. With $L = F(q || \tau)$ and parameters in q (and possibly τ), natural-gradient flows implement

$\dot{q} = - I(q)^{-1} \nabla_q F$, and similarly for τ,

which (i) respects statistical geometry, (ii) improves alignment under distribution shift, and (iii) composes cleanly across membranes when I is block-structured.

4. Links to other optimizers.
    - Mirror descent with KL as the Bregman divergence yields the same direction as a natural gradient step.

- In exponential families, I(θ) equals the Hessian of the log-partition, enabling efficient second-order-like behavior without full Hessians.

Takeaway. KL does triple duty: it gives us (a) a tractable upper bound (via F), (b) a coding ledger (accuracy–complexity), and (c) a *metric* (Fisher) that turns "move toward τ" into well-conditioned, invariant steps. This is the geometric reason VFE/KL are reliable as the one scalar for distance, code, and control in Logostics.

3. Where They Co-Emerge (Survey)

3.1 Neuroscience & active inference (perception, learning, action)

Core move. The brain is modeled as a generative process τ with internal states q that minimize *free energy* F(q||τ).

- Perception: update q(z|x) to reduce instantaneous F (predictive coding = prediction errors weighted by precision).
- Learning: slowly tune τ's parameters to increase model evidence (decrease long-run F).
- Action: choose policies that minimize expected free energy G = E_future[F], which decomposes into epistemic (information gain) + pragmatic (goal realization) terms.
- Precision control: dynamics hinge on *precision* (inverse noise). Misallocated precision overweights/underweights errors, changing the descent field on F—our membrane view treats these as adjustable boundary gains.
- Markov blankets: factorization at sensory–active boundaries makes F additive; message passing implements local KL contractions that sum to a global decrease.

Borrow for Logostics. Use (i) instantaneous F for fast perceptual contraction, (ii) expected F for policy selection, (iii) per-membrane precision as tunable boundary gains.

3.2 Machine learning: VI, VAEs, amortization, reparameterization

Core move. Variational inference maximizes ELBO = −F.

- Amortization: learn an encoder $q_\phi(z|x)$ to approximate posteriors for all x—alignment pressure flows through a single network (cheap inference, global contraction).
- Reparameterization trick: write $z = g_\phi(\varepsilon, x)$, $\varepsilon \sim N(0,I)$, to get low-variance gradients of ELBO/F—crucial for stable, scalable descent on F.

- Accuracy–complexity control: β-VAE and KL annealing explicitly budget the KL term KL(q||p_τ(z)) to trade reconstruction vs parsimony; bits-back coding interprets F as codelength.
- Hierarchies & flows: hierarchical VAEs, normalizing flows, and diffusion-ELBOs refine the same objective (tighter posteriors, better geometry).
- PAC-Bayes link: ELBO/F bounds generalization; minimizing F aligns with minimizing expected description length (MDL) and PAC-Bayes complexity.

Borrow for Logostics. Use amortized q for fast inner contractions; schedule β (or α-KL) per membrane to shape where complexity is paid; prefer reparameterizable families to keep gradients well-behaved.

3.3 Control as inference & KL-control; linearly-solvable MDPs

Core move. Cast control as probabilistic inference: optimal behavior is the distribution over trajectories that (a) accrues reward (negative energy) while (b) stays close—in KL—to passive dynamics.

- KL-regularized objectives: minimize E[cost] + λ KL(π || π_o) or, equivalently, maximize soft value functions (entropy-regularized RL).
- LMDPs / path-integral control: with certain dynamics/costs, the Bellman equation becomes linear after an exponential transform; policies arise from *Boltzmannizing* costs (a free-energy form).
- Soft Q/actor-critic: "temperature" λ sets a precision-like gain; small λ → exploit, large λ → explore (epistemic probing emerges when combined with model uncertainty).
- Robustness: the KL term bounds how far controlled dynamics deviate from a trusted baseline—exactly the membrane-safety intuition.

Borrow for Logostics. Treat policy selection as minimizing expected F with a KL trust region around τ's passive dynamics; tune λ as a boundary precision; use linearly-solvable structure where possible for tractable planning.

3.4 Robotics: generalized gradient descent on VFE

Core move. Replace bespoke estimators/controllers with a unified VFE descent in *belief space*.

- State estimation: beliefs q over latent pose/maps are updated by ∇_q F; sensor likelihoods define accuracy, priors/dynamics define complexity.
- Action/perception loop: actions reduce expected F by seeking observations that disambiguate beliefs (active sensing) while pursuing goals encoded in τ.
- Generalized coordinates: modeling trajectories (positions, velocities, accelerations) extends the blanket; priors become smoothness/physics; controllers become "move to minimize future F."
- Embodiment: proprioceptive/exteroceptive membranes split F into local terms—motor reflex arcs implement steepest-descent moves without global solves.
- Sim-to-real: KL-regularized adaptation keeps policies near reliable passive models, reducing sim-to-real drift.

Borrow for Logostics. Operate entirely in belief space; expose per-sensor/per-actuator F budgets; implement natural-gradient steps for better conditioning; enforce KL trust regions for safe adaptation.

3.5 Computational psychiatry & theoretical biology: organisms minimizing VFE

Core move (psychiatry). Symptoms are framed as *precision misallocation* and prior mis-specification in the free-energy loop.

- Anxiety & hallucination spectra: overweighted priors (high prior precision) vs overweighted sensory noise (low likelihood precision) tilt updates, changing the F landscape and trapping q away from τ.
- Interoception: bodily signals form inner membranes; mis-calibrated interoceptive precision skews valuation and action (expected-F term).
- Treatment as retuning: pharmacology/therapy shift precision or reshape priors, effectively reparametrizing τ or rebalancing membrane budgets.

Core move (theoretical biology). Life = maintaining non-equilibrium steady states by minimizing long-run free energy.

- Homeostasis/allostasis: organisms act to keep states in low-surprise regions defined by τ (niche priors).

- Morphogenesis: developmental programs appear as priors; growth follows F-descent subject to physical constraints; cellular membranes are literal Markov membranes.
- Chemotaxis & foraging: policies that maximize evidence of being "in the right place" are just expected-F minimizers.

Borrow for Logostics. Model value alignment as *niche priors over* τ; instrument per-membrane precision to detect maladaptive regimes; use expected-F policies to recover healthy attractors (homeostatic q→τ).

Section takeaway. Across brains, ML, control, robots, and living systems, the same duet—VFE + KL—plays three roles at once: (1) tractable inference (distance), (2) disciplined compression (code), and (3) robust decision-making (control). Logostics inherits this as a portable design law: minimize F locally at each membrane, bound KL moves globally, and let expected-F shape actions that make tomorrow's alignment easier than today's.

4. From Survey to Synthesis: Logostics q(t) → τ(t)

4.1 Treat $F[q \parallel \tau]$ as a global contraction functional

Definition. With generative law $\tau$and agent state q, the (variational) free energy

$$F[q \parallel \tau] \;=\; \mathrm{KL}\left(q(z \mid x) \;\parallel\; p_\tau(z \mid x)\right) \;-\; \log\, p_\tau(x)$$

is our global scalar of alignment.

Contraction claim (moving target). Let q_t follow the descent field $\dot{q}_t = -\nabla_q F[q_t \parallel \tau_t]$. If $F$is L-smooth in $q$and $\tau_t$drifts with speed bounded by $\parallel \dot{\tau}_t \parallel \leq \delta$, then along trajectories:

$$\frac{d}{dt} F[q_t \parallel \tau_t] \;\leq\; -\;\parallel \nabla_q F \parallel^2 \;+\; C\,\delta,$$

for some Cdepending on $\partial F / \partial \tau$. Hence, when $\tau$is slow (δ small), Fdecreases on average and acts as a Lyapunov-like functional: $q(t) \to \tau(t)$ in the sense of shrinking KL.

Discrete two-loop routine (operational):

1. Inner loop (inference): $q \leftarrow q - \eta_q \nabla_q F$ until ΔF plateaus.
2. Outer loop (model/goal): $\tau \leftarrow \tau - \eta_\tau \nabla_\tau F$ on slower timescale (or via held-out evidence).

3. Monitor: treat F as the single global KPI; stop if F contracts while its variance across boundaries (Sec. 4.2) remains bounded.

Why this works. Fupper-bounds surprise, equals expected codelength, and induces a well-conditioned geometry for updates (Sec. 2). Using it as *the* scalar makes alignment measurable, optimizable, and auditable.

4.2 Markov membranes: localizing alignment via conditional independencies

Membrane idea. Suppose τfactorizes across a Markov membrane (blanket) into conditionally independent parts $\{k\}$:

$$p_\tau(x, z) \;=\; \prod_k \quad p_{\tau_k}(x_k \mid z_k)\, p_{\tau_k}(z_k)$$with sparse couplings.

Then the free energy adds:

$$F[q \parallel \tau] \;=\; \sum_k \quad (\; -\; \mathbb{E}_{q_k}\log\, p_{\tau_k}(x_k \mid z_k) \;+\; \mathrm{KL}(q_k \parallel p_{\tau_k}(z_k))) \;+\; \text{couplings},$$
$$\underbrace{\qquad\qquad}_{F_k}$$

where couplings vanish or are small when the membrane approximates conditional independencies well.

Local budgets and diagnostics.

- Per-membrane budgets: enforce $\mathrm{KL}(q_k \parallel p_{\tau_k}) \le B_k$ with Lagrange multipliers λ_k; this yields *β-schedules per boundary* (high-noise sensors get looser budgets; brittle actuators get tighter ones).
- Precision tuning: each membrane carries its precision (gain) on prediction errors; rebalancing these gains steers where contraction pressure is applied.
- Health checks: plot $\{F_k\}$and $\{\mathrm{KL}_k\}$ time-series; divergence isolated to a few k flags boundary mis-specification rather than global model failure.

Message passing (practical). Because $F = \sum_k \quad F_k$, you can update $q_k$by local gradients using only messages through the blanket (expected sufficient statistics). This yields scalable, asynchronous q→τ updates.

4.3 Instantaneous vs expected free energy for policy selection

Instantaneous free energy.

$$F_t \equiv F[q_t \parallel \tau_t]$$

contracts *current* beliefs given the *current* model. Use it for perception/learning on fast timescales.

Expected free energy (planning). For a candidate policy πover horizon H,

$$G(\pi) = \mathbb{E}_{q(o_{1:H},s_{1:H}|\pi)}[-\log p_\tau(o_{1:H} \mid C) + \mathrm{KL}(q(s_{1:H} \mid o_{1:H},\pi) \parallel p_\tau(s_{1:H} \mid o_{1:H}))],$$

where Cencodes preferences/goals. The two terms read as:

- Pragmatic: $-\log p_\tau(o_{1:H} \mid C)$— seek preferred outcomes (low surprise under C).
- Epistemic: the KL term — seek information that aligns future posteriors with the model (resolve ambiguity).

Policy rule (receding horizon).

1. Roll out a small policy set $\{\pi\}$with learned dynamics; estimate $\hat{G}(\pi)$ by particles or analytic surrogates.
2. Choose $\pi^{\backslash *} \in \arg\min_\pi \hat{G}(\pi)$; execute its first action; replan (MPC style).
3. Trust region: constrain $\mathrm{KL}(\pi \parallel \pi_{\text{baseline}}) \leq \epsilon$ for safety and stability (keeps behavior inside a membrane of reliable dynamics).

When to use which.

- F (instantaneous): belief updates, fast contraction, diagnosis.
- G (expected): action selection, exploration–exploitation without ad-hoc bonuses, safe adaptation via KL trust regions.

4.4 Natural-gradient flows and invariance under reparameterization

Local geometry. KL induces the Fisher metric $I(\theta)$. The natural gradient is

$$\widetilde{\nabla}_\theta F \;=\; I(\theta)^{-1}\,\nabla_\theta F,$$

giving steps that are *invariant* to how parameters are written and that follow the intrinsic curvature of the statistical manifold.

Block-natural flows (membrane-aware). If $q$and τdecompose across membranes, the Fisher blocks are near block-diagonal:

$$I(\theta) \;\approx\; \mathrm{diag}(I_1, \dots, I_K).$$

Update each block independently:

$$\theta_k \leftarrow \theta_k - \eta_k\, I_k^{-1}\,\nabla_{\theta_k} F_k,$$

with occasional cross-block corrections when couplings matter. This preserves scalability and stability.

Trust regions = KL constraints. Equivalently, solve

$$\min_{\Delta\theta} \;\; \nabla_\theta F^\top \Delta\theta \text{s.t.} \Delta\theta^\top I(\theta)\,\Delta\theta \le \varepsilon,$$

whose solution is a natural-gradient step with step size set by ε. This matches mirror descent with KL as the Bregman divergence and connects to TRPO/NPG in RL.

Computational recipe (fast and clean).

- Use reparameterization for low-variance ∇F.
- Approximate $I^{-1}$with K-FAC/Gauss–Newton or Fisher-vector products (CG).
- Schedule per-membrane step sizes $\eta_k$and KL caps ε_k to keep local moves safe while maintaining global progress.

Practical guarantee (empirical). Compared to vanilla gradients, natural-gradient flows on Freduce sensitivity to ill-conditioning, improve convergence under distribution shift, and remain robust when $\tau$is updated concurrently—exactly the regime of q(t) → τ(t).

Section takeaway. Make F the single global KPI; factor it by membranes to localize pressure; use instantaneous $F$to contract beliefs and expected G to choose actions; and move by natural gradients under KL trust regions. This turns Logostics into a concrete control recipe for aligning $q(t)$ to a moving $\tau(t)$ with stability, scalability, and principled exploration.

5. Design Pattern: Building q → τ Systems

5.1 Two-timescale routine: inner KL contraction, outer τ retuning

Goal. Make "alignment" a repeatable control loop: a fast inner loop that contracts $q$toward $\tau$by minimizing instantaneous free energy $F[q \parallel \tau]$, and a slower outer loop that reshapes τ(parameters, structure, or preferences) to increase evidence and stabilize contraction.

Inner loop (fast; milliseconds–seconds in software, cycles in hardware).

- State update. $q \leftarrow q - \eta_q \nabla_q F[q \parallel \tau]$(use reparameterized, low-variance gradients).
- Natural step (preferred). $q \leftarrow q - \eta_q \, I(q)^{-1} \nabla_q F$, with $I$the Fisher metric (blockwise per membrane).
- KL trust region. Enforce $\mathrm{KL}(q_{\text{new}} \parallel q_{\text{old}}) \leq \epsilon_q$ to prevent overshoot.
- Precision tuning. Adapt per-membrane error gains (likelihood/prior precisions) to keep gradient magnitudes comparable.

Outer loop (slow; seconds–minutes, or per episode/epoch).

- Model retuning. $\tau \leftarrow \tau - \eta_\tau \nabla_\tau F[q \parallel \tau]$ on held-out or rolling windows; include structure moves (e.g., adding a latent, pruning a factor).
- Preference retuning (for control). Update goal priors $C$by minimizing expected free energy Gunder validation rollouts.
- Budget rebalance. Reallocate KL budgets B_k across membranes to equalize marginal contraction yields (Sec. 5.2).

Scheduling heuristic.

- Run N_{inner}inner steps per outer step; start with $N_{\text{inner}} \in [5,20]$.
- Increase N_{inner} when ΔF per outer step becomes noise-like; decrease when drift in τinvalidates inner equilibria.

Stop conditions.

- F decreases and coefficient of variation over a window falls below γ.
- Per-membrane KLs stabilize within budgets B_k.
- No boundary shows persistent positive local Lyapunov (Sec. 5.2).

5.2 Diagnostics: per-membrane KLs, contraction indices, escape-time maps

Per-membrane accounting. With a Markov membrane factorization, track for each boundary k:

$$F_k \;=\; -\,\mathbb{E}_{q_k}\log\, p_{\tau_k}(x_k \mid z_k) \;+\; \mathrm{KL}\,(q_k(z_k \mid x_k) \;\|\; p_{\tau_k}(z_k)).$$

Log time series of:

- Accuracy $\quad_k$ and Complexity $\quad_k$ (the split of F_k).
- Precision $\quad_k$ (estimated inverse noise or error gain).
- Grad-norm $\quad_k$ ($\| \nabla_{q_k} F_k \|$) as a stress indicator.

Contraction index (global and local).

- Instantaneous: $\chi_t \equiv -\frac{d}{dt} F_t \approx -\frac{F_t - F_{t-\Delta}}{\Delta}$.
- Normalized: $\hat{\chi}_t = \chi_t / (F_{t-\Delta} + \varepsilon)$.
- Local (per membrane): same definitions for $F_{k,t}$.
  Interpretation: positive χ= contracting; negative = diverging. Maintain moving averages $\bar{\chi}^{(W)}$ to reject noise.

KL-efficiency (where are bits best spent?).

$$\mathrm{Eff}_k \;\equiv\; \frac{-\Delta F_k}{\Delta\,\mathrm{KL}_k + \varepsilon}$$

Higher Eff $\quad_k$ means "one extra bit of complexity buys more contraction" at membrane k. Reallocate budgets $B_k$toward high-Eff $\quad_k$ boundaries.

Escape-time maps (boundary-surfing cartography).

- Idea. Probe a grid of boundary hyperparameters (e.g., precisions, β_k, step sizes) and record time to contraction T_{esc} (iterations needed for $F$to cross a target reduction).

- Procedure.
    1. Choose a 2-D slice of hyperparameters (α, β).
    2. For each grid point, run the inner loop from a fixed q_0; stop when F falls by ρ or a cap is reached.
    3. Plot $T_{\text{esc}}(\alpha, \beta)$; mark non-convergent regions.
- Use. Identify ridges/valleys corresponding to alignment phase transitions; set operating points near valleys for speed or on ridge shoulders for creativity (Sec. 5.3).

Bifurcation sensors.

- Sign-flip detector. Flag when $\bar{\chi_k}^{(W)}$ crosses zero with significant slope.
- Local Lyapunov proxy. $\lambda_k \approx \frac{1}{W}\sum_t \log\frac{\|\delta q_{k,t+1}\|}{\|\delta q_{k,t}\|}$ using two close trajectories. Positive λ_k indicates local chaos/misalignment pressure.

Dashboards (minimum viable).

- Top row: F, $\hat{\chi}$, variance across k.
- Middle rows: per-membrane F_k, Accuracy $_k$, Complexity $_k$, Eff $_k$.
- Bottom: escape-time map for current 2-D slice; markers for last 10 runs.

5.3 Boundary surfing: controlled excursions at alignment phase flips

Rationale. Novel structure often appears where contraction becomes marginal—near phase flips in alignment dynamics. We want exploratory excursions that harvest insight without losing global stability.

Detect the surf-zone.

- A membrane k is in surf-zone when:
  (i) $\bar{\chi_k}^{(W)} \approx 0$(marginal contraction),
  (ii) Eff $_k$ peaks relative to neighbors, and
  (iii) escape-time contours show steep gradients (near a ridge).

Open the boundary (controlled).

- Temporarily relax constraints on k: raise budget $B_k \to B_k(1+\delta_B)$, lower β_k, or widen the policy trust region $\epsilon_{\pi,k}$.

- Simultaneously tighten global safeguards: decrease other B_j, shrink global step size, and cap $\Delta F > 0$excursions by $\Delta F_{\max}$.

Exploration schedule (receding-horizon).

1. Tag a surf-membrane $k^{\setminus *}$.
2. Run H inner-loop steps with relaxed constraints on $k^{\setminus *}$; select actions by minimizing expected free energy Gbut with larger epistemic weight on $k^{\setminus *}$ (inflate its ambiguity term).
3. Checkpoint every $h$steps: if global Fexceeds cap or $\bar{\chi}^{(W)} < -\delta$, rollback to last checkpoint (store q, τ, RNG seeds).
4. Consolidate. If excursion yields net F decrease after return to nominal budgets, adopt new structure (e.g., keep added latent, adjust precision priors) and update escape-time map around the discovered point.

Safety invariants.

- Global KL cap. $\sum_k \quad \Delta\mathrm{KL}_k \le K_{\max}$over any window W.
- Baseline proximity. $\mathrm{KL}(\pi \parallel \pi_0) \le \epsilon$ (policy cannot drift too far from trusted dynamics).
- Monotone-on-average. Require $\sum_{t=t_0}^{t_0+W} \quad \Delta F_t \le 0$ even if short spikes occur.

Heuristics that work in practice.

- One-at-a-time surfing. Never open more than one boundary per window W.
- Short excursions. Keep H small (e.g., 5–20 steps); insight tends to emerge quickly near ridges.
- Elastic budgets. Use proportional-integral control:
  $B_k \leftarrow B_k + \kappa_p(\text{target Eff} - \text{Eff}_k) + \kappa_i \sum \text{error}$.

What counts as “insight”?

- Structural gain. A proposal (new latent, new factor link, modified prior) that reduces baseline F after returning to nominal budgets.
- Compression gain. Persistent drop in Complexity $_k$ without loss of Accuracy $_k$.
- Policy gain. Lower expected Gon held-out rollouts with unchanged caps.

Putting it together (playbook).

1. Stabilize. Run two-timescale routine to steady contraction.
2. Map. Generate escape-time maps around current operating point.
3. Surf. Select one membrane with peak Eff and marginal χ; open it under safety caps; roll forward H steps.
4. Evaluate. Accept only if global Fimproves and invariants hold; otherwise rollback.
5. Record. Update maps and budgets; move to the next membrane or terminate.

Section takeaway. Treat alignment as steady contraction punctuated by brief, guarded excursions. The two-timescale routine keeps you stable; per-membrane diagnostics tell you *where* to push; boundary surfing tells you *how* to push—just enough to cross a ridge and come back with a better model.

6. When KL Is Not Enough

6.1 α-divergences to trade off mode-seeking vs mode-covering

Why: $\mathrm{KL}(q \parallel p)$ is *mode-seeking*: it heavily penalizes putting mass where p has none, which can collapse to one mode under multi-modal $p_\tau(z \mid x)$. Conversely, $\mathrm{KL}(p \parallel q)$ is *mode-covering*: it penalizes missing support, often spreading $q$too thin. α-divergences interpolate:

$$D_\alpha(p \parallel q) = \frac{1}{\alpha(1-\alpha)}(1 - \int p^\alpha q^{1-\alpha}\, d\mu), \alpha \in \mathbb{R} \setminus \{0,1\}.$$

Limits:

- $\alpha \to 1 \Rightarrow D_\alpha \to \mathrm{KL}(p \parallel q)$ (covering).
- $\alpha \to 0 \Rightarrow D_\alpha \to \mathrm{KL}(q \parallel p)$ (seeking).
- $\alpha = \frac{1}{2}$ yields (scaled) Hellinger; symmetric and robust.

Logostics use: per-membrane, let the inference loss be

$$F_k^{(\alpha)} \;=\; \mathbb{E}_{q_k}[-\log\, p_{\tau_k}(x_k \mid z_k)] \;+\; \beta_k\, D_\alpha\,(q_k(z_k \mid x_k) \;\parallel\; p_{\tau_k}(z_k)).$$

Tune α_k by regime:

- Exploration / novelty: α ↑(toward covering) to avoid premature collapse.
- Exploitation / precision: α ↓(toward seeking) to commit to sharp explanations.

Diagnostics: if per-membrane *missed-mode alarms* (e.g., multi-modal residuals, persistent "ghost" predictions) fire, increase $\alpha_k$or mix in a covering term (see 6.3).

6.2 Wasserstein for support mismatch and dynamics in data space

Why: When $q$and p_τ have disjoint or weakly overlapping supports, KL and α-divergences become unstable or uninformative. Optimal transport distances compare distributions *via movement in data space*:

$$W_c(p,q) = \inf_{\gamma \in \Gamma(p,q)} \mathbb{E}_{(u,v)\sim\gamma}[\, c(u,v)\,],$$

with cost c (e.g., $c(u,v) = \| u - v \|_2$). For $c = \|\cdot\|$in $\mathbb{R}^d$, W_1 has the Kantorovich–Rubinstein dual used in WGANs.

Advantages:

- Geometry aware: penalizes *how far* mass must move to align; meaningful with non-overlap.
- Curriculum friendly: provides informative gradients earlier in training or after domain shifts.

Logostics use (two patterns):

1. Decoder-space transport. Add a transport term on reconstructions:

$$F^{\text{hyb}} = F + \lambda_W\, W_1 (p_\tau(x \mid z \sim q), p_{\text{data}}(x)),$$

estimated with a 1-Lipschitz critic. Helpful when sensors (membranes) undergo distribution shifts; the critic becomes a *boundary critic*.

2. Latent transport for structure moves. When adding a new latent factor or remapping components, align *prior* aggregates by W_2 (closed forms exist for Gaussians), avoiding KL blow-ups during re-initialization.

When to prefer Wasserstein: severe covariate shift, sharp manifold-to-manifold gaps (e.g., domain adaptation), or early training of powerful decoders where KL gives vanishing/unstable signals.

6.3 Hybrid objectives and schedule switching across training phases

Motivation: No single divergence suits all phases. Use hybrids and schedules to balance stability (OT), parsimony (KL/α), and sample-efficient exploration (expected free energy).

A. Additive hybrid (membrane-aware)

Per membrane k:

$$\mathcal{L}_k = \underbrace{-\mathbb{E}_{q_k}\log\, p_{\tau_k}(x_k \mid z_k)}_{\text{accuracy}} \;+\; \underbrace{\beta_k\, D_{\alpha_k}(q_k \,\|\, p_{\tau_k}(z_k))}_{\text{complexity}} \;+\; \underbrace{\lambda_{W,k}\;\; W_{1,k}}_{\text{transport}}\;.$$

Tuning rules:

- If Eff $_k$ (contraction per KL bit) is low but WGAN critic has strong signal, increase $\lambda_{W,k}$.
- If Complexity $_k$ rises without accuracy gain, decrease β_k or raise α_k to cover modes.
- If transport critic saturates (Lipschitz penalty dominates), lower $\lambda_{W,k}$ and return pressure to KL/α.

B. Phase schedule (typical)

1. Warm-start (geometry finding): High λ_W, mid α(0.3–0.7), small β. Goal: align coarse supports.
2. Commit (parsimony): Reduce $\lambda_W \to 0$; anneal $\alpha \downarrow 0$ (toward KL(q||p)); raise $\beta \uparrow$to compress.
3. Refine / explore (boundary surfing): Temporarily bump $\alpha \uparrow$ on the target membrane $k^{\backslash *}$and allow small $\lambda_{W,k^{\backslash *}} > 0$; keep global KL caps (Sec. 5.3).
4. Stabilize: Lock $\alpha = 0$, $\lambda_W = 0$; optimize pure $F$with natural gradients.

C. Policy coupling (expected free energy)

During phases 1–3, select actions by minimizing expected F, but inflate epistemic weight for membranes carrying $\lambda_{W,k} > 0$or $\alpha_k > 0.5$. This drives queries that reduce support mismatch (collect data where the OT critic is most informative).

D. Safe switching criteria

Switch or roll back when any holds:

- Monotone-on-average violated: windowed sum $\sum \Delta F > 0$.
- Global KL exceeds cap: $\sum_k \quad \Delta \mathrm{KL}_k > K_{\max}$.
- Critic pathology: OT critic gradient norms explode or collapse for $> m$steps.

E. Minimal pseudocode

```
for phase in [WARM, COMMIT, SURF, STABILIZE]:
  set (alpha_k, beta_k, lambdaW_k) by phase
  while not converged_in_phase:
    # inner loop
    q <- q - eta_q * natgrad_q( sum_k L_k )
    if phase in [WARM, SURF]: update OT critics per membrane k with GP penalty
    # outer loop
    tau <- tau - eta_tau * natgrad_tau( sum_k L_k )
    # policy
    a_t <- argmin_pi  E_pi[ expected_free_energy(pi; weighted_epistemic_by_k) ]
    # safety & diagnostics
    enforce_KL_caps(); check_monotone_on_average(); adapt per-membrane budgets
    if violation: rollback_or_phase_backoff()
```

Takeaway. Use α-divergences to shape *what kind* of alignment you want (seek vs cover), Wasserstein to get meaningful signals under support mismatch, and hybrid, membrane-aware schedules to move from coarse support alignment to parsimonious contraction under F. Keep the guardrails from Sec. 5: per-membrane budgets, KL caps, expected-F planning, and monotone-on-average descent.

7. Case Studies (Sketches)

7.1 Active inference agent with expected-F policies

Setting. 2D target-seeking agent with ambiguous sensors (noisy range + bearing) and simple dynamics s_{t+1} = f(s_t, a_t) + noise. Preferences C favor being near a goal region.

Generative law tau. p_tau(o_t | s_t) Gaussian sensors; p_tau(s_{t+1} | s_t, a_t) local linearization; prior over trajectories encodes smooth motion. Internal state q carries a belief over s_t (mean mu_t, covariance Sigma_t) and a policy posterior over short horizons.

Objective.

- Instantaneous free energy: F_t = accuracy_t + complexity_t for current o_t.
- Planning objective (expected free energy): for policy pi over horizon H,
  G(pi) = E_pi[pragmatic_cost + epistemic_term],
  where pragmatic_cost = -log p_tau(o_{1:H} in C), epistemic_term approx mutual information I(s;o).

Algorithm (receding horizon).

1. Perception: update (mu_t, Sigma_t) by minimizing F_t via natural gradient in belief space.
2. Rollouts: sample K policies pi^k (action sequences); for each, simulate H steps using q-linearized tau; evaluate G(pi^k) using analytic epistemic approximations (variance reduction) plus goal likelihood.
3. Trust region: select pi* = argmin G subject to KL(pi || pi_baseline) <= eps.
4. Execute first action of pi*, observe o_{t+1}, repeat.

Diagnostics. Track F_t, G(pi*), per-sensor F_k, precision gains, and policy KL to baseline. A healthy loop shows decreasing average F, bounded policy KL, and rising epistemic gain early in episodes.

Outcome sketch. Compared to a myopic controller, the agent actively seeks disambiguating views (epistemic term), reaches the goal with fewer steps, and remains stable under sensor dropout by tightening the policy KL cap.

7.2 VAE with membrane-wise KL budgeting

Setting. Multimodal VAE for image+audio pairs (x^img, x^aud). Each modality is a membrane k in {img, aud}; we want aligned factors but different KL budgets per membrane due to SNR differences.

Model.

- Encoders q_phi^k(z_k | x^k), factorized latents z = (z_img, z_aud) with a small shared subspace z_sh and private subspaces z^{prv}_k.
- Decoder p_tau^k(x^k | z_k, z_sh). Priors p_tau(z_sh), p_tau(z^{prv}_k).

Loss (per minibatch).
For membrane k:
L_k = -E_q log p_tau^k(x^k | z_k, z_sh) + beta_k KL(q_phi^k(z_k | x^k) || p_tau(z_k)).
Global:
L = sum_k L_k + beta_sh KL(q_phi(z_sh | x^img, x^aud) || p_tau(z_sh)).

Budgeting rule.

- Choose targets B_img, B_aud, B_sh (in nats per sample).
- Adaptive multipliers:
  beta_k <- beta_k * exp(eta_beta * (KL_k / B_k - 1)).
  This keeps each membrane's complexity near its budget, preventing the clean modality from carrying the noisy one.

Training loop.

1. Reparameterized gradients for encoders/decoders to minimize L.
2. Every M steps update beta multipliers from observed KLs.
3. Optional warm-start with Wasserstein penalty on reconstructions if supports are misaligned (lambda_W > 0, Sec. 6.2), then anneal lambda_W -> 0.

Diagnostics. Per-membrane Accuracy_k, Complexity_k, Eff_k = (-Delta L_k) / (Delta KL_k + eps). Plot KL_sh vs KL_prv to ensure the shared factor captures genuinely cross-modal structure.

Outcome sketch. Membrane-wise budgeting yields cleaner shared latents, better cross-modal retrieval, and robustness when one modality degrades; overall free energy drops smoothly without one encoder collapsing.

7.3 KL-control for safe exploration under model uncertainty

Setting. Unknown dynamics with a trusted baseline model m_0 and a learned model m. We want exploration that improves m while staying close to safe behavior.

Dynamics and policies.

- Passive dynamics under baseline: p_0(s_{t+1} | s_t) and default policy pi_0(a_t | s_t).
- Controlled policy pi(a_t | s_t) aims to reduce long-run cost J while respecting KL trust to pi_0.

KL-regularized control objective.
Minimize over pi:
J(pi) = E_pi[sum_t c(s_t, a_t)] + lambda * sum_t KL(pi(·|s_t) || pi_0(·|s_t)),
with c combining task loss and an epistemic proxy (variance or disagreement between m and an ensemble m^{(e)}).

Connection to free energy.

- The KL term is the control-as-inference regularizer; lambda acts like precision.
- If we lift to trajectory space and Boltzmannize costs, J is a path free-energy. Minimizing J aligns with minimizing expected F when preferences and model uncertainty are encoded in tau.

Algorithm (safe exploration loop).

1. Uncertainty estimation: train an ensemble of dynamics models; let u_t be predicted variance or ensemble disagreement at (s_t, a_t).
2. Define c(s,a) = task_cost(s,a) - k_ep * u_t (negative cost for informative actions) with caps to avoid perverse incentives.
3. Solve for pi using soft actor-critic or linearly-solvable MDP techniques, enforcing statewise KL(pi || pi_0) <= eps via dual variables.
4. Execute with monitoring: if rolling KL or realized cost breaches caps, project pi toward pi_0 (increase dual lambda); log data and refit ensemble.

Diagnostics. Plot (i) statewise policy KLs, (ii) cumulative cost vs baseline, (iii) model error on a validation buffer, (iv) expected free energy estimates under pi. Success = steady reduction in model error and expected F while cumulative cost remains within safety margins.

Outcome sketch. Compared to unregularized exploration, KL-control gathers informative data in high-uncertainty regions without large deviations from safe behavior, preserves stability across

regime shifts, and accelerates convergence of tau while keeping the system within trusted membranes.

8. Theoretical Notes

8.1 Conditions for F as a Lyapunov-like function with moving τ

Setting. Let x be observed, $q_t \equiv q(\, z \mid x; \theta_t\,)$ the agent state, and τ_t the (possibly time-varying) generative law. Write

$$F_t \equiv F[q_t \parallel \tau_t] = \mathrm{KL}(q_t \parallel p_{\tau_t}(z \mid x)) - \log\, p_{\tau_t}(x).$$

Dynamics. Consider the joint flow

$$\dot{\theta}_t = -G_q^{-1}(\theta_t)\, \nabla_\theta F[q_t \parallel \tau_t], \dot{\phi}_t = -G_\tau^{-1}(\phi_t)\, \nabla_\phi F[q_t \parallel \tau_t],$$

where ϕ parameterizes τ and G_q, G_τ are positive-definite preconditioners (e.g., Fisher blocks; Sec. 4.4).

Sufficient conditions for average decrease of F. There exist $m, L > 0$, $B \geq 0$ such that:

1. *(Strong convexity in q)* For fixed τ, $F(\cdot \parallel \tau)$ is m-strongly convex in the coordinates of q induced by G_q:

$$\langle \Delta q,\, \nabla_q^2 F\, \Delta q \rangle \;\geq\; m \parallel \Delta q \parallel_{G_q}^2.$$

2. *(Lipschitz gradients)* $\parallel \nabla_q F(q \parallel \tau) - \nabla_q F(q' \parallel \tau) \parallel_{G_q^{-1}} \leq L \parallel q - q' \parallel_{G_q}$, and similarly for τ under G_τ.

3. *(Bounded τ drift)* $\parallel \frac{d}{dt}\, p_{\tau_t}(z \mid x) \parallel_{*} \leq B$ for a norm dual to the one used for KL's Gateaux derivative (intuitively, τ moves slowly).

Then along the flow:

$$\frac{d}{dt} F_t = -\parallel \nabla_q F_t \parallel_{G_q^{-1}}^2 - \parallel \nabla_\tau F_t \parallel_{G_\tau^{-1}}^2 + \underbrace{\langle \partial_\tau F_t,\, \dot{\tau}_t \rangle}_{\text{target drift}} \;\leq\; -\, m \parallel q - q^{*}(\tau_t) \parallel_{G_q}^2 + C\, B,$$

for some $C > 0$. Hence, if $B < C^{-1} m\,\epsilon$, the *windowed* decrease holds:

$$\frac{1}{T}\int_{t}^{t+T} \frac{d}{ds} F_s \, ds \;\leq\; -\epsilon \text{ for large enough } T,$$

i.e., F is Lyapunov-like on average for a slowly moving τ_t. Discrete steps inherit this with step sizes η_q, η_τ small enough ($\eta \leq 2/(L+m)$) and standard stochastic-approximation conditions.

Remarks.

- Using natural gradients $G_q = I(q), G_\tau = I(\tau)$ enlarges $m$and shrinks L(better conditioning).
- With membrane factorization, strong convexity/Lipschitz hold blockwise; global guarantees follow from block-diagonal dominance plus bounded cross-terms.
- The nuisance $-\log p_{\tau_t}(x)$ is independent of q; it only contributes via target drift (bounded by (3)).

8.2 Relation to MDL, PAC-Bayes, and stochastic thermodynamics

MDL (coding). The expected message length under a bits-back scheme equals F:

$$L(q;\tau) = \mathbb{E}_q\,[-\log\, p_\tau(x \mid z)] + \mathrm{KL}\,(q(z \mid x) \parallel p_\tau(z)) = F[q \parallel \tau].$$

Minimizing $F$is two-part code minimization: accuracy (data fit) + complexity (posterior–prior divergence). This is the MDL rationale for our accuracy–complexity budgeting.

PAC-Bayes (generalization). For predictors $f$drawn from a *posterior* Q and a *prior* P, a typical PAC-Bayes bound states (with probability $1-\delta$):

$$\mathrm{Risk}(Q) \;\leq\; \widehat{\mathrm{Risk}}(Q) + \sqrt{\frac{\mathrm{KL}(Q \parallel P) + \log\,(1/\delta)}{2n}},$$

where n is the sample size. If we identify Q with $q(\,\cdot \mid x\,)$ (or a parametric posterior over τ's parameters) and P with p_τ, the complexity term in F is exactly the capacity penalty in PAC-Bayes. Thus, low Fimplies good compression *and* tight generalization control.

Stochastic thermodynamics (non-equilibrium free energy). Writing the non-equilibrium free energy of beliefs over microstates z:

$$\mathcal{F}(q) \;=\; \underbrace{\mathbb{E}_q[E_\tau(z)]}_{\text{internal energy}} - \underbrace{H(q)}_{\text{entropy}} \;=\; \mathrm{KL}(q(z) \,\|\, p_\tau^{\mathrm{eq}}(z)) \;-\; \log\, Z_\tau,$$

reveals the same KL structure (with $p_\tau^{\mathrm{eq}} \propto e^{-E_\tau}$). Under Markovian dynamics with detailed balance, free-energy dissipation gives

$$\frac{d}{dt}\,\mathrm{KL}(q_t \,\|\, p_\tau^{\mathrm{eq}}) = -\,\sigma_t \;\le\; 0,$$

where $\sigma_t$is the entropy-production rate. Interpreting $p_\tau(\,z \mid x\,)$ as a *target* measure, our descent on Fmirrors relaxation to equilibrium; with a moving τ_t, the extra work term corresponds to the "housekeeping heat" needed to track a shifting environment. This is the thermodynamic reading of q(t) → τ(t): alignment is *free-energy dissipation* with bounded driving.

8.3 Links to score matching and diffusion objectives

Score matching (Fisher divergence). For an unnormalized model $p_\theta(x) \propto e^{-E_\theta(x)}$, Hyvärinen's score matching minimizes

$$\mathcal{J}_{\mathrm{SM}}(\theta) = \mathbb{E}_{p_{\mathrm{data}}}[\,\| \nabla_x \log\, p_\theta(x) \,\|^2 + 2\,\Delta_x \log\, p_\theta(x)],$$

which equals (up to constants) the Fisher divergence $\mathbb{E}\,\| \nabla_x \log\, p_\theta - \nabla_x \log\, p_{\mathrm{data}} \,\|^2$. For small Gaussian corruption, denoising score matching (DSM) is equivalent to maximizing a variational lower bound on $\log\, p_\theta(x)$. Hence, optimizing the *score* can be seen as minimizing a local surrogate of KL, with the metric given by data-space gradients rather than latent-space posteriors.

Diffusion models. Let the *forward* noising SDE be

$$dx_t = f(x_t, t)\, dt + g(t)\, dW_t, t \in [0,1],$$

and the *reverse* generative SDE use the score $s_\theta(x,t) \approx \nabla_x \log\, p_t(x)$. The standard training objective is DSM over time:

$$\min_\theta \;\; \mathbb{E}_{t\sim\mathcal{U},\, x_t\sim p_t}[\;\; \| \, s_\theta(x_t,t) - \nabla_x \log\, p_t(x_t) \,\|^2 \;\;].$$

Two connections to our F:

1. Path-space KL. Under mild conditions, training a perfect score makes the reverse process match the forward-noised data in path law, i.e., minimizes a KL on trajectories $\mathrm{KL}(\mathbb{P}_\theta \parallel \mathbb{P}_{\mathrm{data}})$. Thus, diffusion training is a *score-space route* to KL contraction—but in data/path space rather than latent space.
2. ELBO view for variational diffusion. Variants (e.g., variational diffusion, consistency models with likelihood) derive explicit ELBOs whose negative equals a free-energy like functional across time steps. In that light, diffusion models minimize a sum of local free energies, each a quadratic (Fisher-metric) approximation to a KL discrepancy at time t.

Bridging recipes (Logostics).

- When latent $q(\,z \mid x\,)$ is hard to fit but data-space geometry is informative (images, audio), hybridize: optimize $F[q \parallel \tau]$ in latent space while adding a score penalty or OT term in data space (Sec. 6.2–6.3). The score term supplies gradients when supports are far apart; F supplies parsimony and calibrated uncertainty.
- Use Tweedie's formula to translate denoisers into posterior mean estimators; this ties DSM heads to the accuracy part of ELBO while the KL part regularizes latents.
- For time-varying τ_t(nonstationary data), view diffusion noise levels as a *curriculum*: early steps behave like Wasserstein/score matching (good signals with disjoint supports), late steps merge into KL/ELBO-dominated refinement.

Takeaway. Score matching and diffusion do *local, geometry-aware* alignment in data/path space, while VFE/KL do *global, code-aware* alignment in latent/model space. In Logostics, we treat them as complements: scores/OT get us *onto* the right manifold; F(and its Fisher geometry) keeps us *on* it with principled compression, uncertainty, and control.

9. Open Problems and Next Steps

9.1 Membrane discovery and adaptive partitions

Problem. We assumed membranes (conditional-independence partitions) were known. In practice they must be discovered, validated, and adapted as tasks, sensors, and goals change.

Research directions

- Structure learning via free-energy gain. Define a search over factorizations $\mathcal{M}$ and pick $\mathcal{M}^{\setminus *}$that maximizes *net contraction per parameter*:

$$\mathcal{M}^{\backslash *} = \arg\max_{\mathcal{M}} \frac{\Delta F_{\text{global}}(\mathcal{M})}{\Delta\#\text{params}(\mathcal{M}) + \lambda \Delta\text{latency}(\mathcal{M})}.$$

- Information bottleneck at the blanket. Learn minimal boundary variables B minimizing

$$\min_{p(B|X)} I(B; X_{\text{inside}}) + I(B; X_{\text{outside}}) - \beta\, I(B; X_{\text{cross}}),$$

to keep only cross-coupling essentials.

- Causal audits. Use interventional tests (do-operations on candidate boundary nodes) and retain those whose removal increases F or breaks Granger/transfer-entropy links across sides.
- Adaptive splitting/merging. Maintain a pool of membranes; split when a boundary's residuals become multi-modal or when Eff $_k$ drops persistently; merge when cross-membrane KL falls below a tolerance.
- Neural implementations. Mixture-of-experts with learned routers trained on $-\Delta F$ as a routing reward; routers implicitly implement membrane assignment.

Benchmarks & evaluation

- Synthetic worlds with known partitions (Sensor fusion; multimodal VAEs).
- Real: robotics stacks (vision–proprioception–haptics), multimodal clinical data.
- Metrics: global F, per-membrane stability, routing sparsity, latency, and recovery under sensor failure.

9.2 Multi-agent Logostics and shared τfields

Problem. How do many agents align to a shared or partially shared τ? What are the equilibria and failure modes?

Research directions

- Shared latent graph. Let agents i have private q_i and *overlapping priors* $p_{\tau^{\text{shared}}}(z_s)$. Optimize

$$\min_{\{q_i\},\, \tau^{\text{shared}}} \sum_i F[q_i \,\|\, \tau_i] + \gamma\, \mathrm{KL}(\bar{q}_s \,\|\, p_{\tau^{\text{shared}}}),$$

where $\bar{q}_s$ aggregates shared posteriors; γtunes consensus strength.

- Communication as variational messages. Messages are *sufficient statistics* of boundary posteriors; bandwidth/prioritization decided by marginal ΔF per transmitted nat.
- Game-theoretic layer. Add costs for divergence between agents' policies and a social baseline: $\sum_i \quad \lambda \mathrm{KL}(\pi_i \parallel \pi_{\mathrm{soc}})$. Study Nash vs team-optimal equilibria in expected-F games.
- Heterogeneous confidence. Agents carry precision priors (credences). Consensus via precision-weighted averaging of shared factors; adversarial robustness by capping any agent's influence through a trust-region in KL.
- Cultural drift & curriculum. Evolve τ^{shared} with federated ELBO: each agent contributes gradients scaled by local contraction indices; schedule updates to prevent rapid swings.

Evaluation

- Cooperative control (multi-robot exploration) with partial observability.
- Human–AI teams: measure global F, sample efficiency, safety violations, and fairness (variance of per-agent Fimprovements).

9.3 Toward principled boundary metrics for "insight regions"

Problem. We used heuristics (Eff, χ, escape-time maps) to find creative "ridges" where new structure emerges. We need theory-grounded metrics and tests.

Candidate metrics

- Second-variation curvature at the blanket.

$$\kappa_k \;=\; \lambda_{\max}(\nabla^2_{q_k} F_k) \;-\; \lambda_{\min}(\nabla^2_{q_k} F_k).$$

High spectral spread suggests stiff/flat mixed directions—fertile for new factors.

- Fisher–Ruppeiner mismatch. Compare Fisher blocks of inside vs outside models at a boundary:

$$\Delta I_k = \parallel I_k^{\text{in}} - I_k^{\text{out}} \parallel_{\mathrm{op}}.$$

Large mismatch flags *geometry conflict* that a new latent or coupling could resolve.

- Predictive novelty index. Boundary-conditioned predictive entropy drop:

$$\mathrm{PNI}_k = H(o_{\text{cross}} \mid B_k) - H(o_{\text{cross}} \mid B_k, \text{probe}).$$

Probes are small experimental actions; large drops imply high epistemic leverage.

- Phase-flip score. A smoothed sign-change detector on local contraction:

$$\Phi_k = \Pr\left(\bar{\chi}_k^{(W)} \text{ crosses } 0 \text{ with slope } > s_0\right).$$

- Transport stress. Dual OT critic value at the boundary $W_{1,k}$: high transport with stagnant KL indicates support mismatch—a cue to restructure.

Decision rules (opening a boundary)
Open membrane $k$for surfing when a composite score

$$S_k = a_1\,\mathrm{Eff}_k + a_2\,\Phi_k + a_3\,\mathrm{PNI}_k + a_4\,\Delta I_k + a_5\,W_{1,k}$$

exceeds a threshold, with $a_i$learned from past wins (successful insights) via bandit or Bayesian optimization.

Validation protocols

- Ablation with rollback. Require that adopting the proposed change (new latent, new coupling, re-precisioning) reduces baseline Fafter returning to nominal budgets.
- Counterfactual replay. Re-run recent data with and without the change; accept only if expected Gimproves on held-out policies.
- False-discovery control. Use Benjamini–Hochberg on batches of boundary proposals to limit spurious “insights.”

Tooling roadmap

- Open-source membrane lab: (i) per-boundary dashboards, (ii) escape-time map generator, (iii) bandit tuner for S_k, (iv) safe rollback with snapshotting of q, τ, RNG seeds.
- Reference tasks with known “insight regions” (synthetic bifurcations, domain shifts, sensor swaps) to standardize comparisons.

Section takeaway. Three frontiers will mature Logostics from method to discipline: (1) discovering and adapting membranes on the fly; (2) extending alignment to many agents that

share (and negotiate) a moving τ; and (3) grounding our "surfing" intuition in principled boundary metrics that predict when and where insight will pay—then testing those predictions with strict rollback-and-verify protocols.

10. Conclusion

10.1 Why VFE/KL are the workhorse of alignment

Across brains, robots, software learners, and living systems, the pair variational free energy (VFE) and KL divergence keeps winning because it does three jobs at once with a single scalar:

- Distance. It upper-bounds surprise and contracts beliefs to the target posterior—even when the generative law $\tau$is only approximate—so it behaves like a Lyapunov-like "distance-to-truth."
- Code. Its accuracy–complexity split equals expected codelength, so every unit of alignment pressure is audited: are we paying for fit (likelihood) or for belief complexity (posterior–prior drift)?
- Control. Its time-extended form, expected free energy, unifies exploration and exploitation: act to both realize preferred outcomes and gather information that will shrink tomorrow's F.

Add to this the geometry (Fisher metric → natural gradients), decomposability (additivity across Markov membranes), and tractability (reparameterized stochastic gradients, conjugacy, message passing), and VFE/KL becomes not just a theory but a portable engineering primitive. When supports mismatch or multi-modal structure matters, we augment it (α-divergences, Wasserstein, diffusion/score terms), but we return to $F$for auditing, contraction, and steering. That is why VFE/KL serve as the workhorse of Logostics: one objective that measures *how far*, tells us *what it costs*, and shows *how to move*.

10.2 A portable recipe for contracting $q(t)$toward $\tau(t)$across domains

Implement Logostics as a disciplined loop—with guardrails—independent of the substrate (neural, robotic, software, organizational):

1) Make F your single KPI.
Track global F and its split (accuracy/complexity). Log per-membrane F_k and KL_k; enforce KL budgets B_k.

2) Run two timescales.
Fast inner loop: contract $q$by descending F(prefer natural gradients; enforce KL trust regions

between steps).
Slow outer loop: retune τ(parameters, structure, preferences) only after inner contraction stabilizes.

3) Use membranes to localize work.
Factor models and updates across boundaries; allocate and adapt β_k (complexity weights) to equalize *marginal contraction per nat* (Eff $_k$).

4) Plan with expected F.
Select actions/policies by minimizing horizon-G (expected free energy) under a policy KL trust region to a safe baseline. Inflate epistemic weight where diagnostic leverage is highest.

5) Add geometry, not guesswork.
Adopt natural-gradient or mirror-descent steps with Fisher preconditioning (K-FAC/CG in practice). This stabilizes learning under reparameterization and shift.

6) Switch objectives when needed, then come back.
Warm-start with Wasserstein/score-matching when supports are far apart; use α-divergences to trade mode-covering vs mode-seeking during discovery; anneal back to pure $F$for compression and calibrated uncertainty.

7) Instrument the frontier.
Maintain contraction indices, escape-time maps, and composite insight scores to locate alignment phase flips; surf one boundary at a time with rollback and global caps.

8) Enforce safety invariants.
Windowed monotone-on-average decrease in F; caps on cumulative policy KL and per-membrane KL; snapshot/rollback for any violation.

Minimal success test (domain-agnostic).
Over rolling windows: (i) F trends down with bounded variance; (ii) per-membrane KLs respect budgets; (iii) expected Gunder chosen policies decreases on held-out rollouts; (iv) after exploratory boundary openings, baseline F is lower once constraints are restored.

Follow these eight rules and you have a portable alignment routine: steady contraction punctuated by guarded exploration, audited by code length, shaped by geometry, and stabilized by membranes. That's Logostics in practice—q(t) → τ(t) as an implementable law, not a slogan.

References

Amari, S. (1998). Natural gradient works efficiently in learning. *Neural Computation, 10*(2), 251–276.

Arjovsky, M., Chintala, S., & Bottou, L. (2017). Wasserstein GAN. *Proceedings of ICML*.

Beal, M. J. (2003). *Variational Algorithms for Approximate Bayesian Inference* (PhD thesis). University College London.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.

Buckley, C. L., Kim, C. S., McGregor, S., & Seth, A. K. (2017). The free energy principle for action and perception: A mathematical review. *Journal of Mathematical Psychology, 81*, 55–79.

Catoni, O. (2007). *PAC-Bayesian Supervised Classification: The Thermodynamics of Statistical Learning*. IMS.

Cover, T. M., & Thomas, J. A. (2006). *Elements of Information Theory* (2nd ed.). Wiley.

Cuturi, M. (2013). Sinkhorn distances: Lightspeed computation of optimal transport. *NIPS*, 2292–2300.

Dziugaite, G. K., & Roy, D. M. (2017). Computing nonvacuous generalization bounds for deep (stochastic) neural networks. *ICLR*.

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience, 11*(2), 127–138.

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., & Pezzulo, G. (2017). Active inference: A process theory. *Neural Computation, 29*(1), 1–49.

Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. *Journal of Physiology-Paris, 100*(1–3), 70–87.

Friston, K., Parr, T., & de Vries, B. (2017). The graphical brain: Belief propagation and active inference. *Network Neuroscience, 1*(4), 381–414.

Friston, K., Sengupta, B., & Auletta, G. (2014). Cognitive dynamics: From attractors to active inference. *Proceedings of the IEEE, 102*(4), 427–445.

Haarnoja, T., Zhou, A., Abbeel, P., & Levine, S. (2018). Soft actor-critic: Off-policy maximum entropy deep RL with a stochastic actor. *ICML*.

Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. *NeurIPS*.

Hyvärinen, A. (2005). Estimation of non-normalized statistical models by score matching. *Journal of Machine Learning Research, 6*, 695–709.

Jordan, M. I., Ghahramani, Z., Jaakkola, T., & Saul, L. K. (1999). An introduction to variational methods for graphical models. *Machine Learning, 37*(2), 183–233.

Kappen, H. J., Gómez, V., & Opper, M. (2012). Optimal control as a graphical model inference problem. *Machine Learning, 87*(2), 159–182.

Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. *ICLR*.

Kingma, D. P., & Welling, M. (2014). Auto-encoding variational Bayes. *ICLR*.

Levine, S. (2018). Reinforcement learning and control as probabilistic inference: Tutorial and review. *arXiv:1805.00909*.

Martens, J., & Grosse, R. (2015). Optimizing neural networks with Kronecker-factored approximate curvature. *ICML*.

McAllester, D. A. (1999). Some PAC-Bayesian theorems. *Machine Learning, 37*(3), 355–363.

Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective*. MIT Press.

Nguyen, X., Wainwright, M. J., & Jordan, M. I. (2010). Estimating divergence functionals and the likelihood ratio by convex risk minimization. *IEEE Trans. Information Theory, 56*(11), 5847–5861. (α-divergence connections)

Parr, T., & Friston, K. (2019). Generalised free energy and active inference. *Biological Cybernetics, 113*(5–6), 495–513.

Pearl, J. (2009). *Causality* (2nd ed.). Cambridge University Press. (Markov blankets and conditional independence)

Rezende, D. J., Mohamed, S., & Wierstra, D. (2014). Stochastic backpropagation and approximate inference in deep generative models. *ICML*.

Rissanen, J. (1983). A universal prior for integers and estimation by minimum description length. *Annals of Statistics, 11*(2), 416–431.

Rissanen, J. (1989). *Stochastic Complexity in Statistical Inquiry*. World Scientific.

Schulman, J., Levine, S., Abbeel, P., Jordan, M., & Moritz, P. (2015). Trust region policy optimization. *ICML*.

Seifert, U. (2012). Stochastic thermodynamics, fluctuation theorems and molecular machines. *Reports on Progress in Physics, 75*(12), 126001.

Song, Y., & Ermon, S. (2019). Generative modeling by estimating gradients of the data distribution. *NeurIPS*. (Score matching / diffusion)

Tishby, N., Pereira, F. C., & Bialek, W. (2000). The information bottleneck method. *Proceedings of the 37th Allerton Conference*.

Todorov, E. (2007). Linearly-solvable Markov decision problems. *NIPS*, 1369–1376.

van den Berg, R., Hasenclever, L., Teh, Y. W., & Rezende, D. J. (2018). Sylvester normalizing flows. *UAI*. (Posterior flexibility for VI)

Wainwright, M. J., & Jordan, M. I. (2008). Graphical models, exponential families, and variational inference. *Foundations and Trends in Machine Learning, 1*(1–2), 1–305.

Youvan, D. C. (2025, October). *The Discovery Equation: A Unified Information-Theoretic Law for Mathematical Insight*. DOI: 10.13140/RG.2.2.33299.34085.

Zellner, A. (1988). Optimal information processing and Bayes's theorem. *American Statistician, 42*(4), 278–284. (MDL/Bayes coding perspective)

Zhang, T. (2006). From ε-entropy to KL-entropy: Analysis of minimum information complexity density estimation. *Annals of Statistics, 34*(5), 2180–2210.